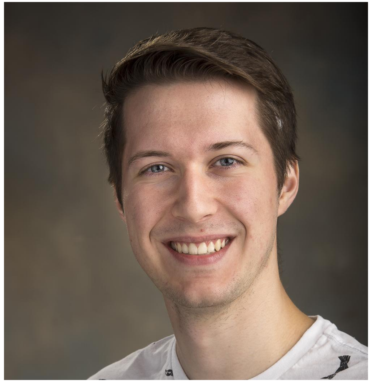
Winning the WAR on Defense

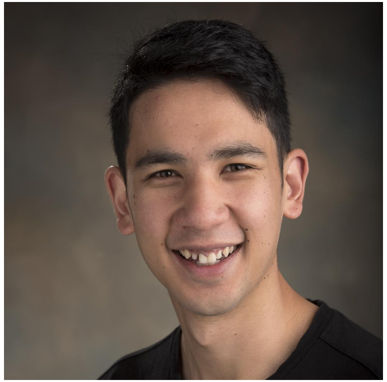
Evaluating the Value Add of defensive players through Sports Info Solutions Data

Collaborators

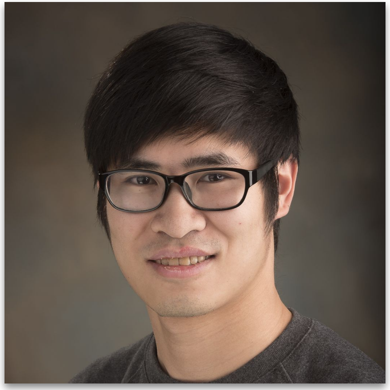


Matthew Reyers
MSc Student / Data Scientist
SFU / Zelus Analytics
matthew_reyers@sfu.ca
@Stats_By_Matt

Meyappan Subbaiah
Data Scientist
Zelus Analytics
meysubb@gmail.com
@msubbaiah1



Dani Chu
Quantitative Analyst
NHL Seattle
danic.stats@gmail.com
@chuurveg



Lucas Wu
PhD Student / Data Scientist
SFU / Zelus Analytics
yifan_wu@sfu.ca
@lucaswu123

What is the most valuable defensive line position and how is value distributed?

The Need

- Counting stats are hard to find on defense
  - Current public data limited to tackle-based counts
- Value is confounded among lineups
  - Limited ability to partition credit
- Teams aren’t limited to offense
  - Evaluation of defense players has large team building implications

The Approach

- Leverage SIS data
    - Pressures, gaps, and line arrangements
- Establish logical play baselines
    - Do separately for passing / rushing
- Calculate value per player per play
- Define summary statistic
    - Additional EPA generated
        - Translate to Wins Above Replacement

Value Added

Value added = Predicted value - Observed value

From public predictive models

Convert each play to expected points based on SIS data

Summary Statistic

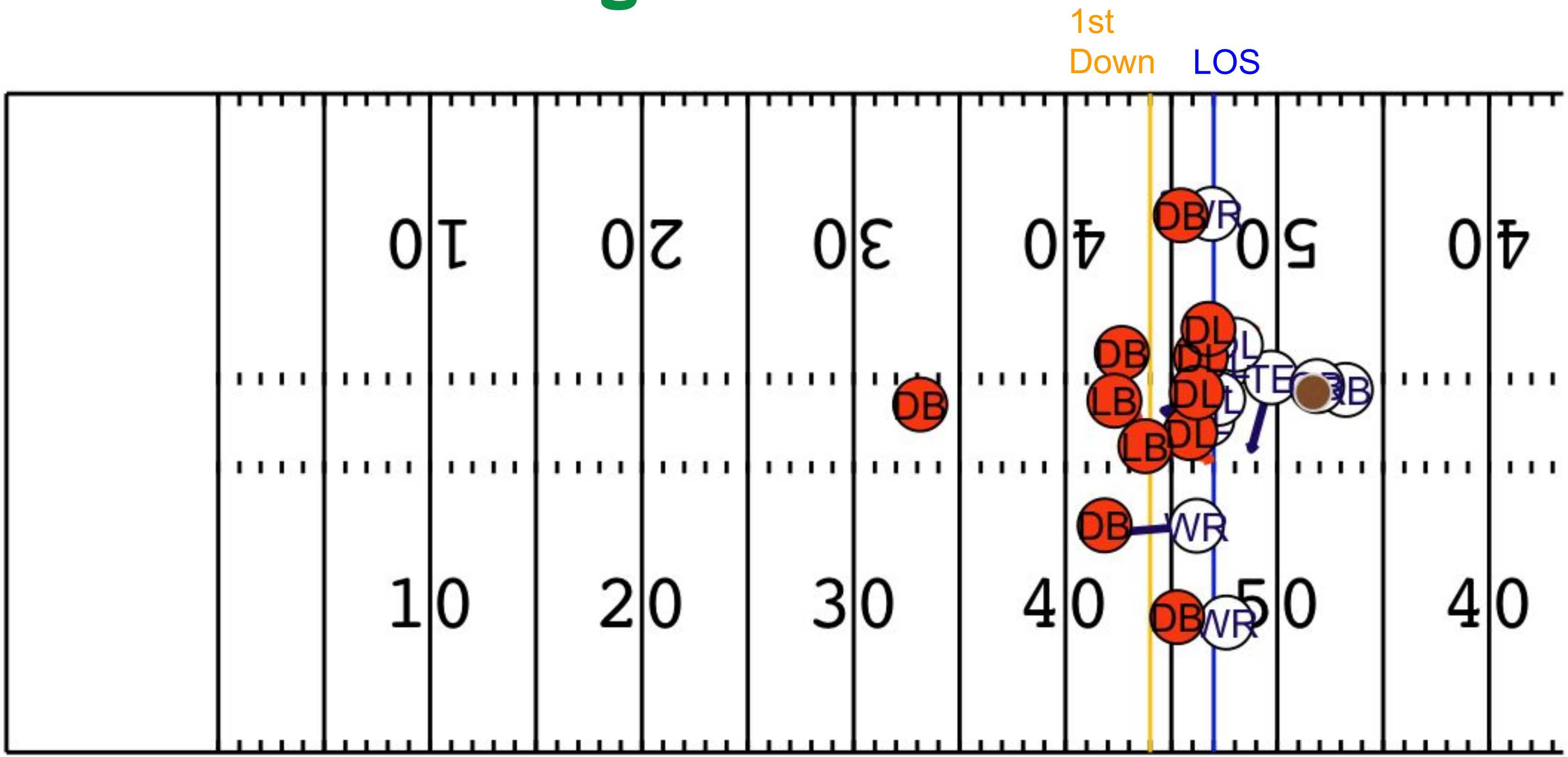
- Value Added in terms of additional EPA generated
- Translate to wins
  - Fit basic linear regression as in nflWAR paper (Yurko, Ventura & Horowitz, 2019)
  - Estimate: 1 Win = 38.40 EPA
- Calculate contributed wins rushing, passing, and combined
  - Per game basis, rushing sample reduced to accommodate Zoo predictions - 1st place solution from 2019-2020 NFL Big Data Bowl Competition for predicting yards gained on given rushing plays

Rushing

The Details: Rushing

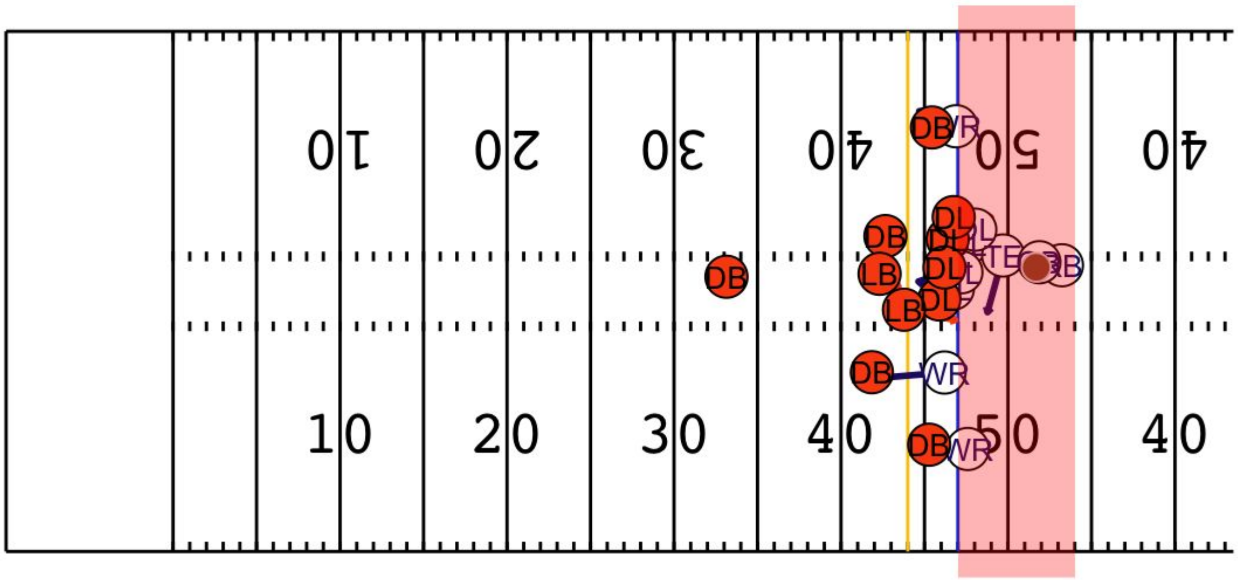
- Baseline
  - Incorporate predictions from the winners of the 2019--2020 NFL Big Data Bowl (the Zoo)
    - Gives predicted yards per play at time of handoff
  - Translate predicted yards to more robust Expected Points Added
- Adjustments
  - Use all rush data to estimate Rusher and Offensive Line effects
  - Adjust predictions

The Details: Rushing



The Details: Rushing

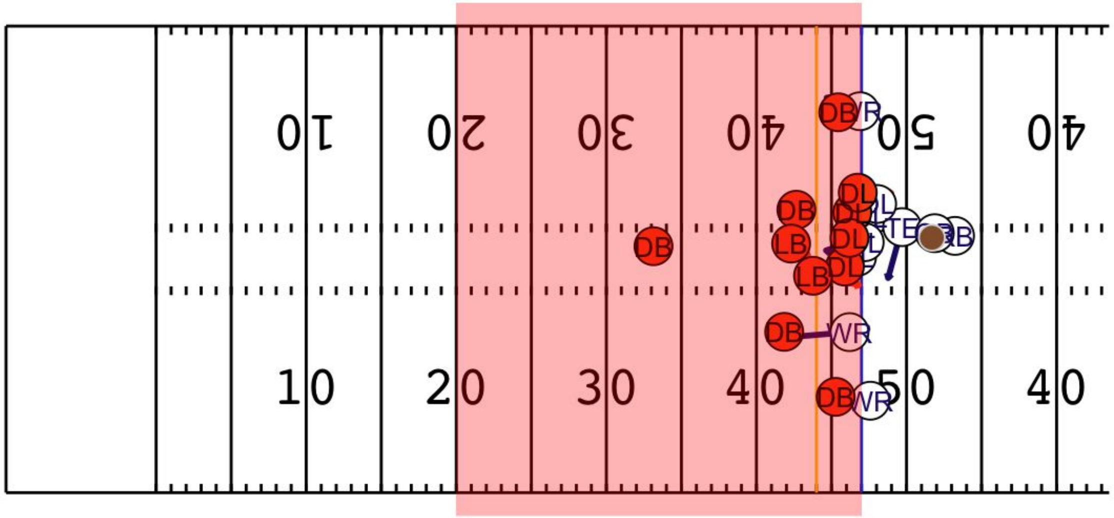
- Value
  - Partition play, 4 outcomes
    - Play ends behind LOS



The Details: Rushing

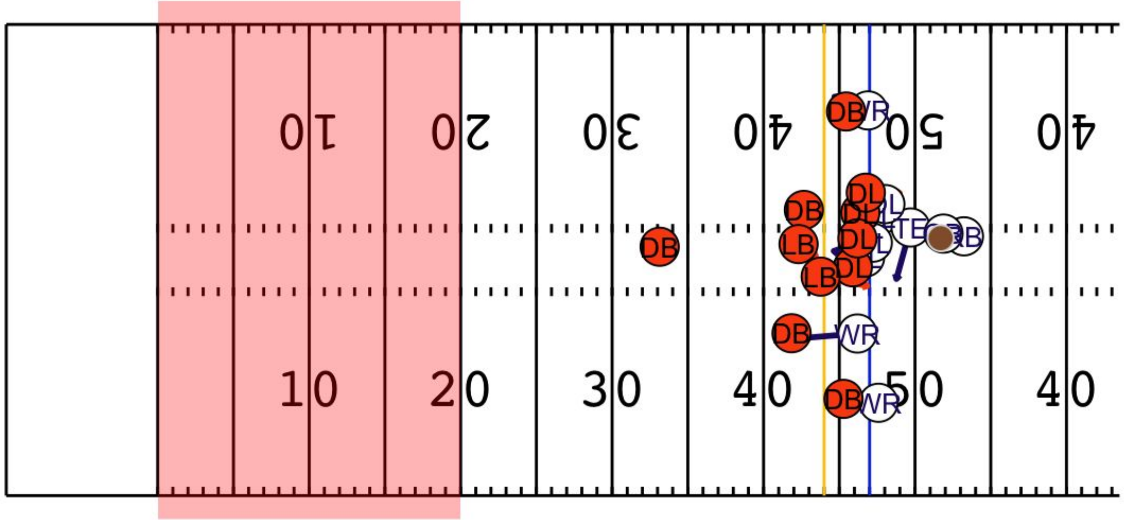
- Value
  - Partition play, 4 outcomes
    - Play ends behind LOS
    - Play ends between LOS and Zoo predicted yards

Predicted yards



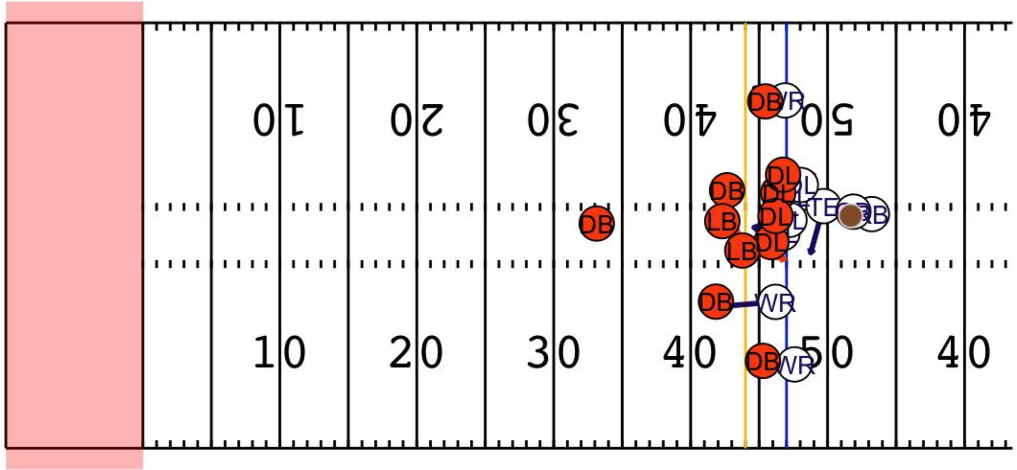
The Details: Rushing

- Value
  - Partition play, 4 outcomes
    - Play ends behind LOS
    - Play ends between LOS and Zoo predicted yards
    - Play ends beyond Zoo predicted yards but is not TD



The Details: Rushing

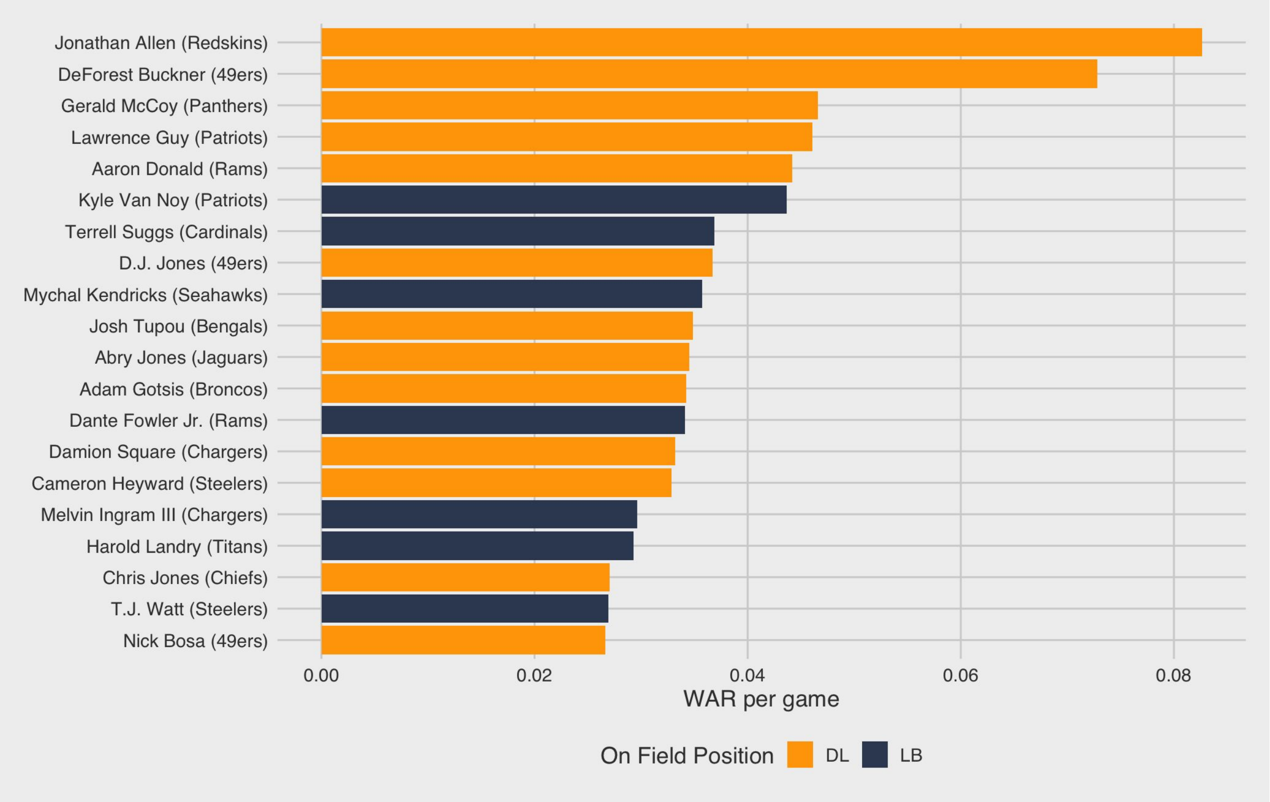
- Value
  - Partition play, 4 outcomes
    - Play ends behind LOS
    - Play ends between LOS and Zoo predicted yards
    - Play ends beyond Zoo predicted yards but is not TD
    - Play is a TD



The Details: Rushing

- Value
  - Partition play, 4 outcomes
    - Play ends behind LOS
    - Play ends between LOS and Zoo predicted yards
    - Play ends beyond Zoo predicted yards but is not TD
    - Play is a TD
- Involvement
  - A player is involved if they record a pressure, tackle, or sack
  - A player gains more credit for being involved, less credit for not

WAR Winners: Rushing (Top 20)



Sample Conclusions

- 🏈 Majority of value on rushing plays accrued by DL
- 🏈 Focus is on Interior Lineman
- 🏈 Larger gap between best DL and average DL than between best LB and average LB

Passing

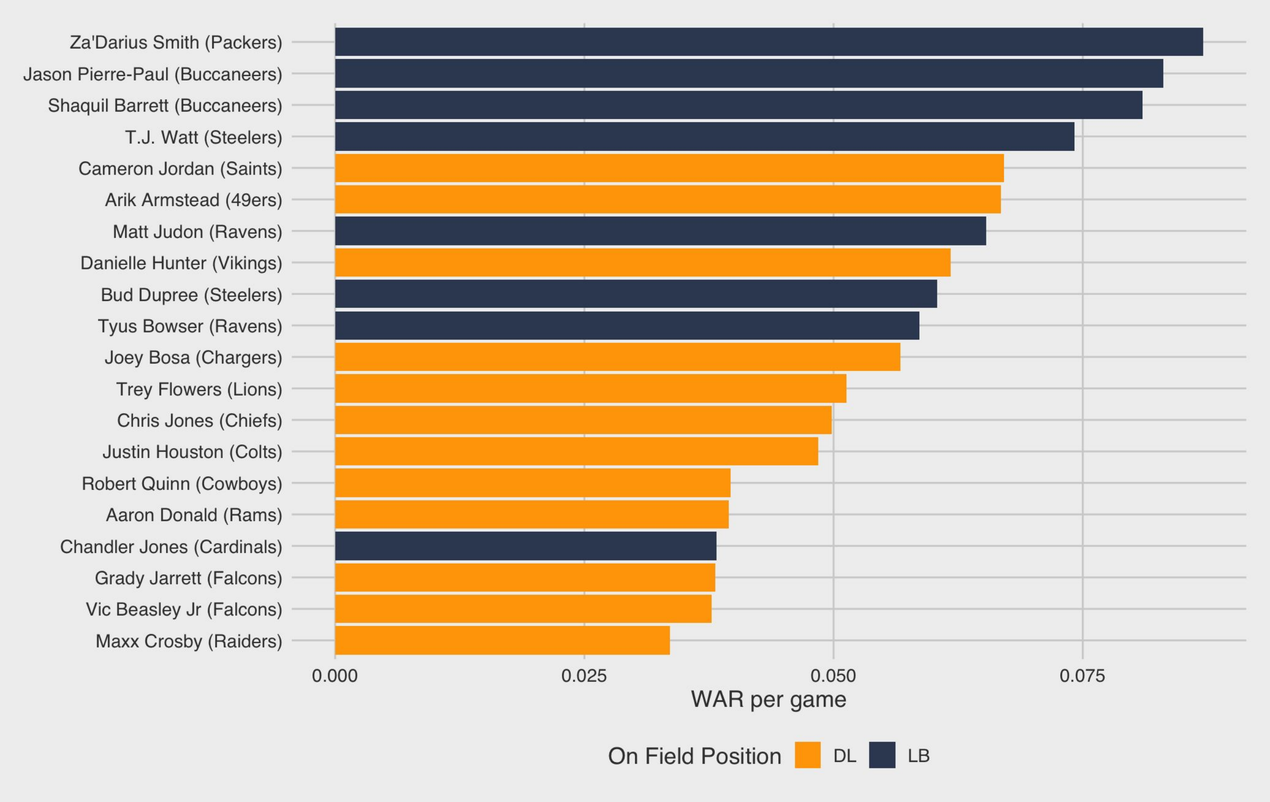
The Details: Passing

- 🏈 Baseline
  - Incorporate completion probability model from nflfastR
    - Estimate two completion probabilities per play
      - If QB pressured
      - If QB not pressured
  - Establish baseline as weighted EPA based on completion probability and target depth
- 🏈 Adjustments
  - Arbitrarily set minimum target depth to 5 yards for value attribution

The Details: Passing

- Value
  - Partition play, 3 outcomes
    - No pressure generated
    - Pressure generated, no sack
    - Sack
- Involvement
  - A player is involved if they generate a pressure, sack, pass breakup, or interception

WAR Winners: Passing (Top 20)



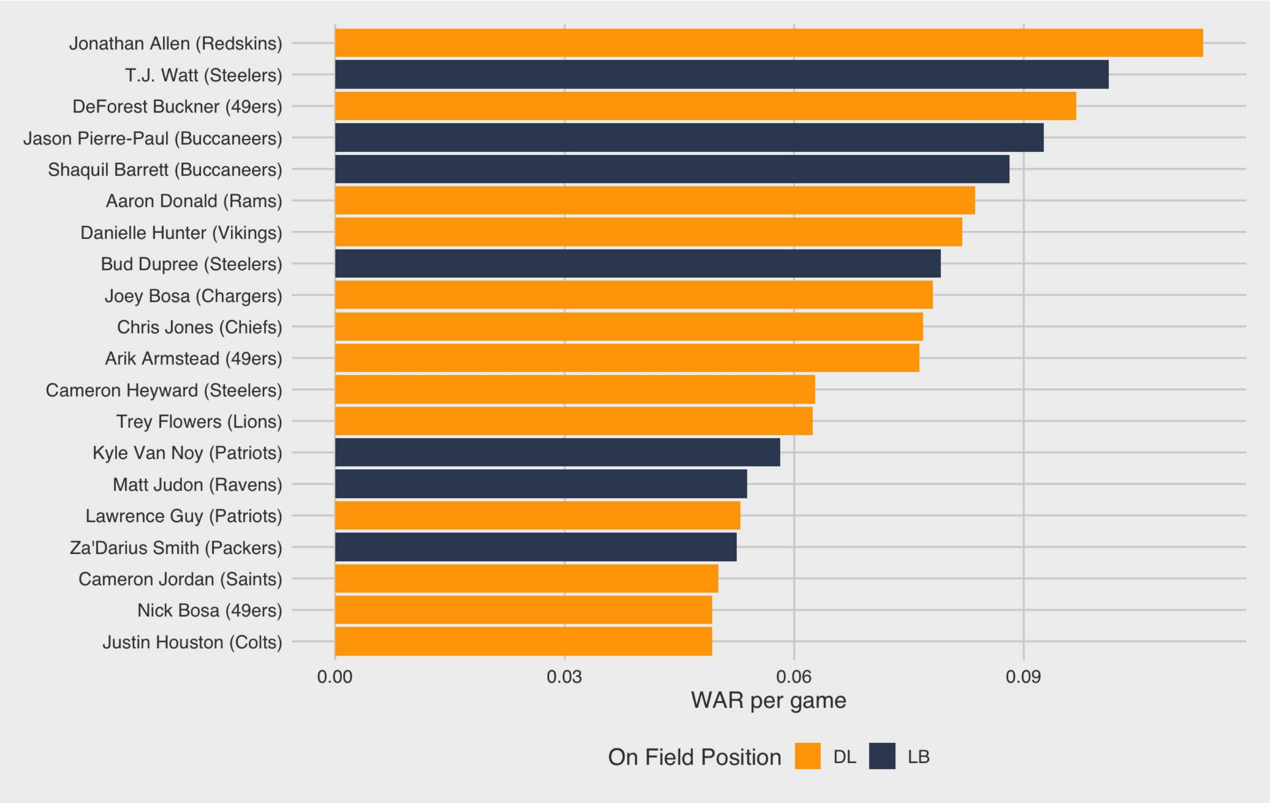
Sample Conclusions

- 🏈 Greater emphasis on Linebackers and Edge Rushers

- 🏈 Value generated by top-tier Linebackers and Edge Rushers is harder to replace than top-tier Interior DL

Rushing and Passing

WAR Winners: Overall (Top 20)



WAR Winners - Top 4 Defensive Teams (DVOA -FO)

WAR Winners for 49ers
WAR values on a per game basis

Player	Team	Cluster Assignment	Roster Position	WAR - Rushing	WAR - Passing	WAR - Overall	Overall Rank
DeForest Buckner		2	DT	0.073	0.024	0.097	3
Arik Armstead		5	DE	0.010	0.067	0.076	11
Nick Bosa		1	DE	0.027	0.023	0.049	19
D.J. Jones		6	DT	0.037	−0.001	0.036	38
Solomon Thomas		5	DE	0.002	−0.024	−0.022	140
Sheldon Day		6	DT	−0.006	−0.017	−0.023	142

WAR Winners for Patriots
WAR values on a per game basis

Player	Team	Cluster Assignment	Roster Position	WAR - Rushing	WAR - Passing	WAR - Overall	Overall Rank
Kyle Van Noy		1	LB	0.044	0.014	0.058	14
Lawrence Guy		5	DE	0.046	0.007	0.053	16
Danny Shelton		6	DT	−0.005	−0.004	−0.009	99
Jamie Collins		1	LB	−0.008	−0.019	−0.027	161
John Simon		1	LB	−0.016	−0.014	−0.030	163

WAR Winners for Ravens
WAR values on a per game basis

Player	Team	Cluster Assignment	Roster Position	WAR - Rushing	WAR - Passing	WAR - Overall	Overall Rank
Matt Judon		1	LB	−0.011	0.065	0.054	15
Tyus Bowser		3	LB	−0.016	0.059	0.043	27
Jaylon Ferguson		1	LB	0.009	0.014	0.023	47
Jihad Ward		3	DE	−0.017	0.008	−0.010	102
Chris Wormley		5	DE	−0.007	−0.011	−0.018	127
Brandon Williams		2	DT	−0.027	−0.029	−0.056	201

WAR Winners for Steelers
WAR values on a per game basis

Player	Team	Cluster Assignment	Roster Position	WAR - Rushing	WAR - Passing	WAR - Overall	Overall Rank
T.J. Watt		1	LB	0.027	0.074	0.101	2
Bud Dupree		1	LB	0.019	0.060	0.079	8
Cameron Heyward		2	DE	0.033	0.030	0.063	12
Javon Hargrave		2	DT	0.014	0.032	0.046	22
Tyson Alualu		5	DE	−0.010	−0.022	−0.032	172

Sample Conclusions

- 🏈 Reasonably balanced mix of DL Interior, Edge, and Linebackers

- 🏈 In our sample, Jonathan Allen, T.J. Watt, and DeForest Buckner were all generating roughly 0.1 WAR per game

Shortcomings of WAR

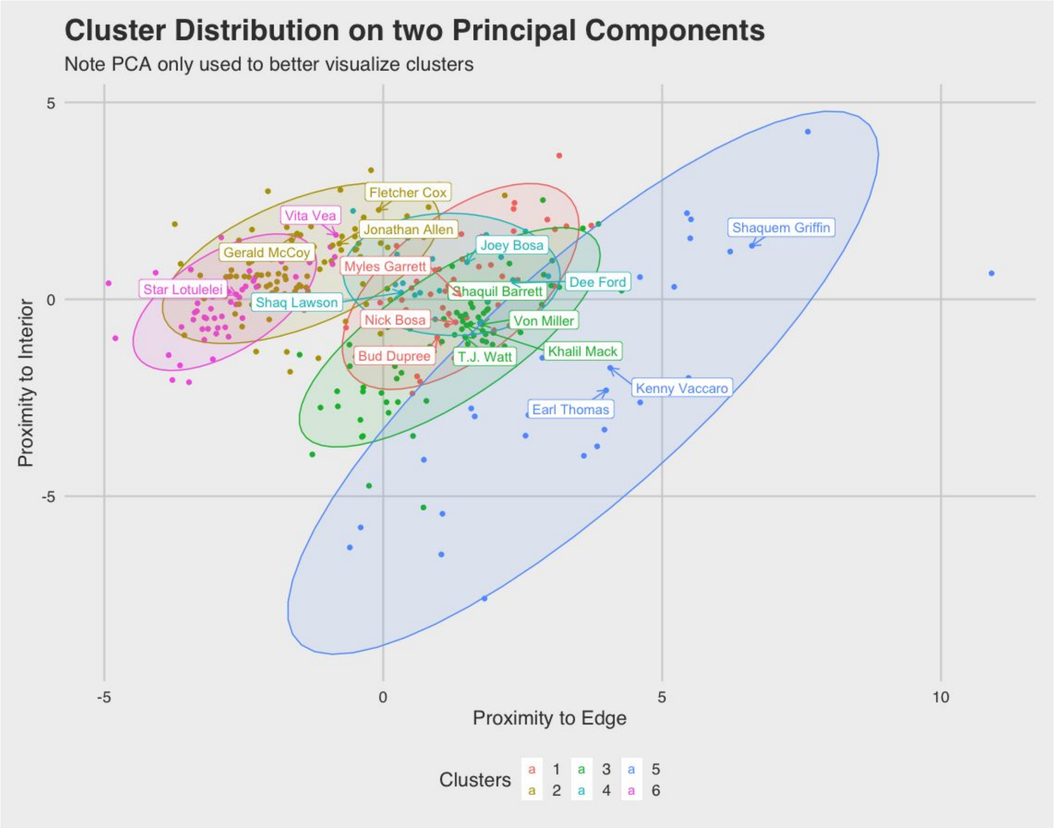
- Rushing sample size is 3 or 4 games per player
  - Zoo predictions only overlapped with first 4 games of SIS data sample
  - Some teams had Byes during this period
- Not large enough sample size overall to conclude with statistical significance
  - We removed a lot of noise but much remains
- Confounding effect on value attribution
  - We report multiple players from the same team in our top players list
  - All talented in their own right though some may be being pulled upwards by exceptional team play

How do the results change if we redefine Defensive Line Positions?

Relabeling Through Clustering

- Incorporate purely usage based statistics
  - Gap assignments, play breakdowns, etc
    - Limit to only Players > 20 snaps
  - Ignore production and salary
    - Goal is to find similarly used players
    - These covariates bias clusters
- Hierarchical clustering
  - Ward's Method - Euclidean Distance
  - 6 clusters

Usage Clusters



Adding Context to the Clusters

Cluster	Key Positive Loadings	Key Negative Loadings	Our Labels	Prototype Player(s)
1	C and D gap usage on right side of line, pass play usage	Short yardage	**Blindside Rushers**	Myles Garrett, Bud Dupree
2	Long yardage, B Gap both sides	Late downs	**Early Down Interior**	Fletcher Cox, Gerald McCoy
3	D gap usage, rushing plays	Interior assignments	**Body on the Edge**	Khalil Mack, Shaquil Barrett
4	C gap usage, passing plays	Early downs	**Multipurpose Outer DL**	Dee Ford, Joey Bosa
5	Short yardage, late downs	Interior assignments	**Edge Rush Package**	Earl Thomas, Solomon Thomas
6	Early downs, A gap, rushing plays	Short yardage	**Early down Nose Tackles**	Star Lotulelei, Vita Vea

Redefining Positional Value

- Use new clusters, explore distribution of Wins Above Replacement
  - Blindside Rushers consistently offer better WAR than other cluster assignments when a replacement player is specified as a generic Defensive Lineman
  - Early Down Interiors are championed by elite talent
    - Beyond the elite talent, group is underwhelming
  - Edge Rush Packages are well distributed with a few standouts
  - Early Down Nose Tackles are the weakest cluster in terms of WAR

In which situations do positional values change?

Situation Dictates Values

- A variety of positions is important to an effective defensive line
- Some valuable situations don’t even call for a player who plays the most valuable position that we have identified
- We would like to determine which groups of positions are most effective for each game situation
    - Use newly created clusters
    - Additional cluster for players with too small a sample to do reasonable inference

Process

- Fit a Bayesian Additive Regression Tree (BART) model in order to ask counterfactual questions
- For example: “What would the outcome of the play look like if we used a different defensive line unit on that play”
- We can measure the effect of each defensive line unit to determine which one would be optimal in each situation

Results - 3rd and 2 rush at the 50 start of 2nd Q

Top 3 Units:

Blindside	Early Interior	Multipurpose Outer DL	Body on the Edge	Edge Rush Package	Early Nose Tackle	Low D-Line Snaps
0	2	0	1	0	1	0
0	2	1	1	0	1	0
1	2	0	0	0	1	0

Despite Blindside being the most valuable position the top two units in this situation do not have a blindside rusher on the line.

Results - 3rd and 10 pass at the 50 start of 2nd Q

Top 3 Units:

Blindside	Early Interior	Multipurpose Outer DL	Body on the Edge	Edge Rush Package	Early Nose Tackle	Low D-Line Snaps
0	1	3	0	0	1	0
0	2	2	0	0	1	0
1	2	1	0	0	1	0

When the offense is likely to pass we see a shift to use more Multipurpose Outer DL than early down interior lineman

Conclusions

- Positional value is heavily dependant on situation
- You do not want all of the players on the defensive line to be the same position
- There is value in variety

Overall Shortcomings and Limitations

- 🏈 Clusters make sense but many players have insufficient sample sizes
  - Two clusters are rather small, many undersampled players would fit these
- 🏈 With the emergence of tracking data, it would be possible to better cluster DL players
  - Movement patterns and trajectories fit a model-based clustering scheme
    - See “Route Identification in the National Football League” for details
- 🏈 WAR values are heavily impacted by a few plays due to brevity of data
  - Analysis would be better handled on whole season worth of data
- 🏈 Baseline predictions for rushing arguably already involve efficacy of players
  - “At time of handoff” means defensive players can influence prediction
  - Downwards pressure on available value for highly talented lines/players
- 🏈 Win Probability omitted from BART model
  - Possible positivity issues 🚩 in specific situations

Wrap-up

- Leveraging SIS data we created
  - An enhanced Wins Above Replacement metric for the often under-analyzed defensive players
  - A new classification scheme for defensive players based on their observed usage
  - A collective measure of value within these new positions, based on observed production
  - An evaluation scheme for causal inference of position groupings for situational considerations

References

Chu, D., Reyers, M., Thomson, J., & Wu, L. Y. (2020). Route identification in the National Football League: An application of model-based curve clustering using the EM algorithm. *Journal of Quantitative Analysis in Sports*, *16*(2), 121-132.

Gordeev, D.1st place solution The Zoo. https://www.kaggle.com/c/nfl-big-data-bowl-2020/discussion/119400. Accessed: 2020-07-19.

Horowitz, M., Yurko, R., and Ventura, S. (2019).nflscrapR: Compiling the NFL play-by-play API for easy use in R. R package version 1.8.3. https://github.com/maksimhorowitz/nflscrapR.

Carl, S., and Baldwin, B. (2020). nflfastR: Functions to Efficiently Scrape NFL Play by Play Data. https://mrcaseb.github.io/nflfastR/.

Schoenfield, D. (2012). “What we Talk about when we Talk about War”. URL http://espn.go.com/blog/sweetspot/post/_/id/27050/what-we-talk-about-when-we-talk-about-war. Accessed: 2020-07-19.

Yurko, R., Ventura, S., & Horowitz, M. (2019). nflWAR: a reproducible method for offensive player evaluation in football. *Journal of Quantitative Analysis in Sports*, *15*(3), 163-183.